DOCUMENT RESUME

ED 155 729 CS 204 217

AUTHOR Vella, Dorothy; Hilgers, Thomas
TITLE An Experiment in Homogeneous Grouping in Freshman Composition Classes.
PUB DATE Aug 77
NOTE 31p.; Report prepared at the University of Hawaii at Manoa

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Ability Grouping; *College Freshmen; Composition (Literary); *Composition Skills (Literary); Educational Research; *English Instruction; *Expository Writing; Grouping (Instructional Purposes); Heterogeneous Grouping; Higher Education; *Homogeneous Grouping; *Writing Skills

ABSTRACT

In order to determine if homogeneous grouping in basic expository writing classes makes teaching more effective, 2,212 incoming college freshmen were given a two-hour essay placement examination and divided into four groups according to the kinds of writing problems they had. They then enrolled in sections designated by placement group, except for certain randomly selected students who were enrolled in seven heterogeneous sections. Near the end of the semester another essay test was administered along with questionnaires to students and faculty members who had participated in the homogeneous groups. The end-of-course essays were rated using the same methods used for rating the pretests. Although the differences between the end-of-course examination scores of the heterogeneous groups and the homogeneous groups were not great enough to be statistically significant, a great majority of both teachers and students were convinced that homogeneous grouping did make teaching in the writing courses more effective. Students in classes of 15 or fewer showed greater improvement than students in larger classes, although once again the difference was not statistically significant. In addition, it was found that students improve in writing proficiency as a result of taking freshman English. (JF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED155729

AN EXPERIMENT IN HOMOGENEOUS GROUPING IN FRESHMAN COMPOSITION CLASSES

by

Dorothy Vella and Thomas Hilgers

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Dorothy Vella
Thomas Hilgers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Department of English
University of Hawaii at Manoa
August 1977

CS204 217

SUMMARY

In 1976-77 the Composition Program of the Department of English at the University of Hawaii at Manoa conducted an experiment in homogeneous grouping in basic expository writing classes (English 100). We gave a two-hour essay placement examination to a total of 2,212 incoming freshmen, and, on the basis of that examination, we divided students into four groups according to the kinds of writing problems they had. In the Fall 1976 and Spring 1977 semesters, we enrolled students in sections designated by placement group--except for certain randomly selected students whom we enrolled in seven heterogeneous sections set up in the spring semester to provide a control group.

Near the end of the spring semester we gave the students in all English 100 classes a second essay exam. We tried to make this posttest as similar as possible to our pretest, but we could schedule the test only during regular class hours and therefore we were forced to make compromises with regard to both the time allotted and the content of the test. To get subjective evaluations to supplement our statistical data, we also distributed two questionnaires: one to all the students in the homogeneous sections, another to all faculty members who had taught the homogeneous sections.

We rated the end-of-course essays written by all students who had taken both the pretest and the posttest, using the same methods--and even the same readers--we had used for the pretest. We then compiled full sets of data on each of these students: ratings on the pretest and the posttest, sex, SAT scores, the type of grouping in the class the student had been in, the size of the class, and the instructor of the class.

The primary question we attempted to answer in analyzing the data was "Does homogeneous grouping in basic expository writing classes make teaching more effective?" The answer was not a clear, emphatic "Yes" or "No."

The statistical data were not conclusive. At all placement levels except Level I, students in the homogeneous groups scored higher on the end-of-course exam than students in the heterogeneous groups. It might be tempting to conclude that homogeneous grouping is more beneficial than heterogeneous grouping for students at the lower levels of proficiency but less beneficial for students at the highest level. But this conclusion cannot be supported by our data because of the compromises we were forced to make in designing the end-of-course exam; the students at Level I might not have been an exception to the general rule if our end-of-course exam had given them greater opportunity to demonstrate the more sophisticated skills they had probably acquired. And, in any case, the differences between the end-of-course exam scores of the heterogeneous groups and the homogeneous groups were not great enough to be statistically significant.

The subjective evaluations of the effects of homogeneous grouping--from both teachers and students--provided a much clearer, more emphatic answer to our primary question than the statistical data did. The great majority of both teachers and students were convinced that homogeneous grouping did make teaching in the basic expository writing courses more effective, did lead to greater improvement in the student's writing proficiency.

The results of our study are not conclusive enough to justify a demand for immediate adoption of homogeneous grouping in our basic expository writing classes. But they certainly do justify further research on the subject.

The need to isolate the effects of homogeneous grouping from the effects of other factors led us into a secondary area of investigation: the effects of class size. We found that students in classes of 15 or fewer students generally showed greater

improvement in writing skills than students in larger classes. The differences were not great enough to be statistically significant. But our data suggest that limiting writing classes to 15 or fewer students would enable students to make greater progress whether the students were grouped homogeneously or not. We did not intend at the outset of this study to investigate the effects of class size, and therefore we made no attempt to distribute our students systematically into classes of various sizes. But our findings certainly seem to justify a study in which this would be done.

We did intend, at the outset of the study, to attempt to answer the questions "Do our students really show improvement in writing proficiency as a result of taking English 100?" and "Do our students meet at least minimum standards of proficiency at the end of the course?" We were pleased to find that our data answered both these questions in the affirmative. The great majority of our students scored higher on the end-of-course exam than they had on the placement exam and in fact, by the end of the course, were communicating at a level that we regarded as meeting at least minimal standards of proficiency.

We certainly are not satisfied with the level of proficiency attained by our students, however. The high reliability of the ratings of the readers of the placement exams indicates that we have established goals. But we are aware that we need to work on clarifying our goals further for ourselves and for our students and we are aware that we need to improve classroom techniques to help students to reach those goals. And this will require further research. This study has been essentially a pilot study; perhaps its greatest value has been that it has suggested important areas which demand further analysis, experimentation, and evaluation.

AN EXPERIMENT IN HOMOGENEOUS GROUPING IN FRESHMAN COMPOSITION CLASSES

RATIONALE AND PROCEDURES

In 1976-77 the Composition Program of the Department of English at the University of Hawaii at Manoa conducted an experiment in homogeneous grouping in basic expository writing classes (English 100, freshman composition). We were concerned--as teachers everywhere were (and still are)--about the decline in verbal scores on the College Entrance Examination Board's Scholastic Aptitude Test (SAT) and about other evidence of a general decline in the writing proficiency of students entering college, and we thought that homogeneous grouping--grouping according to the kinds of writing problems the students had--might help us to increase the effectiveness of our teaching, enable us to give students increased help in developing their writing skills.

We had never given incoming freshmen a placement examination in writing. For students at the upper end of the scale, we had offered an exemption exam that they could volunteer to take if they thought they wrote well enough to be exempted from taking the composition course. For students at the lower end of the scale, we had provided an opportunity to get special help, but this also had been on a volunteer basis. In 1974 we had set up a Writing Workshop, and students who thought they needed individual help in composition could go to one of the Writing Workshop instructors for an interview and, if the instructor gave permission, could enroll in one of the Workshop sections, in which step-by-step help was given. The vast majority of students, however, had not volunteered for either of these options; they had simply waited until registration time and then enrolled in any section of English 100 that was not yet filled and that fitted into their schedules. As a result, there had been a great range in the abilities of the students in our classes: some students were ready to

handle quite complicated, sophisticated writing assignments; others had difficulty writing even one clear, coherent paragraph. This range had made the English 100 sections difficult to teach. If the instructor worked at a level that was suitable for the more proficient students, the less proficient ones found the course too difficult; if the instructor worked at a level that would help students with serious writing problems, the rest were bored.

Every now and then someone had suggested that we give a placement examination--at least to identify the students who should be exempted from the course and the students who needed special help. But everyone realized that giving such an examination to 2,000 to 3,000 freshmen would be a lot of work. Even more important, it would cost a considerable amount of money. And that we had not had. But when, in 1976, some money from the President's Fund for Improvement of Lower Division Instruction became available, the Composition Program asked for--and received--enough money to test all incoming freshmen. So in 1976, for the first time, we required all incoming freshmen to take a placement test in writing proficiency. On the basis of that exam, in the Fall 1976 and Spring 1977 semesters we divided students in English 100 into four groups according to the kinds of writing problems they had. (And for the students at the lowest level we provided special help by assigning tutors and graduate assistants to assist the instructors so that the classes could be broken up into very small groups for work on every stage of the writing process.)

Administering the test for the first time presented many problems of logistics. We wanted to involve the high schools in the process, and so we gave the exam in June at selected high schools--at Baldwin High School on Maui, at Kauai High School on Kauai, at Hilo High School on Hawaii, and at five high schools--Iolani, Farrington, Kailua, Damien, and Aiea--on Oahu. (A list of the dates and places of the exam was sent to all students accepted at Manoa, and they were asked to register to take the test at one of these locations). We gave the test

later at the Manoa campus--twice in August, once in November--for those students who had been accepted after June or who had not been in the islands to take the test in June. Altogether, we tested a total of 2,212 students.

We devised the placement examination ourselves. It was a two-hour examination that required the writing of two essays. Question 1 tested the student's reading comprehension as well as his ability to state ideas clearly and to organize material in a logical fashion: we gave the student two passages on the same subject (in June, the causes of war and the possibility of eliminating war; in August, the legalization of marijuana; in November, the status of women) and asked him to write an essay answering the Question "On what points do the two authors agree and/or disagree?" Question 2 required the student to make a generalization and support it with specific examples or illustrations from his own experience: we gave the student a choice of two one-sentence quotations on widely different subjects (e.g., "Man has become the tool of his tools" and "A way of seeing is also a way of not seeing") and asked him to write an essay defending, attacking, or modifying the sentiment expressed.

Reading the exams was a large undertaking. We chose our readers from among the regular staff of the freshman composition course; only instructors who had been teaching that course were eligible to serve as readers. At the beginning of the reading sessions, to ensure as great a degree of uniformity as possible in our ratings, we spent a good deal of time reading sample papers aloud and discussing their strengths and weaknesses and the relative importance that should be given to those strengths and weaknesses in the total rating. We decided not to assign specific points for specific factors, however; we agreed that the essays should be scored holistically--i.e., read for the total impression they made. We established two levels of exemption--one that carried semester hours' credit, one that did not. The papers that did not earn exemption (the vast majority) we divided into four groups (Group I having the simplest problems, Group IV the

most difficult problems). In deciding to establish four groups (as in deciding to score the essays holistically) we were following a procedure that had proved workable in other such tests (and the procedure that will be followed by the College Entrance Examination Board when it reintroduces its essay test this coming year). Division into four groups helps to ensure reliability in the ratings, for it is relatively easy to decide whether a given paper is in the upper half or lower half of the entire group, and it is relatively easy to decide whether a paper is in the upper or lower portion of that half. Three readers had to agree on the rating of a paper before we considered the decision final. On some papers, the first three readers agreed; on others, there was disagreement and the papers went to as many as six or seven readers before agreement of three was achieved. It took many man-hours of work to complete the reading. But we did the job carefully and conscientiously. And, for the first time, we had information ahead of time about the kinds of problems our incoming freshmen had and how widespread those problems were.

Although the essays were scored holistically, certain factors weighed more heavily than others in the evaluation: the logicality of the thinking and the quality of the evidence presented to support generalizations were naturally given greater weight than mechanical matters such as spelling and punctuation. And, therefore, it is possible to describe the four groups of students in terms of the kinds of writing problems characteristic of each. Group I consisted of students who were judged "ready to go" at a college level, ready to handle college-level writing assignments. The students in Group I read accurately and had good ideas, and in general they had good control of sentence structure. They needed a college-level writing course, but their writing problems were relatively easy ones to deal with: wordiness, pretentious diction, a tendency to stay on a high level of abstraction, some relatively minor weaknesses in organization and coherence, some relatively minor problems of grammar and punctuation. The students who were

placed in Group II had problems similar to those of the Group I students but to a more marked degree. The students placed in Group III had some problems of reading comprehension; many made inaccurate statements about the content of the passages in Question 1. Also, these students had some serious weaknesses in organizing their material, and their control of sentence structure was shaky. They also made quite a few serious errors in grammar and in mechanical matters such as spelling, punctuation, and capitalization. The students placed in Group IV had serious problems of reading comprehension; they made many inaccurate statements about the content of the passages in Question 1--in fact, often completely misunderstood the meaning. Most of them also had serious problems of organization, coherence, sentence structure, grammar, and mechanics. Some students in Group IV had great trouble finding anything at all to say; these students gave extremely incomplete answers, jotting down only a few disconnected sentences in answer to each question.

We sent a notice to each student telling him what group he had been placed in, asking him to bring the notice to registration and sign up for a section designated as belonging to his group. Registration for both spring and fall semesters went smoothly. Before the spring semester, in order to provide a control group for our study of the effects of homogeneous grouping, we set up seven heterogeneous sections (designated as belonging to "Group B"); the names of the students who were asked to enroll in these sections were selected at random by computer from a complete list of students who had taken the placement exam but had not taken English 100 in the fall semester.

From the very beginning of our planning of an evaluation study, we had decided to use a standard pretest-posttest control-group design. And so, soon after the spring semester began, we started to work on plans for the posttest. We naturally wanted our posttest to be as similar as possible to our pretest. But circumstances forced us to make compromises. The students had been given

two hours to complete the placement exam. But we were unable to free a comparable two hours for the end-of-course exam. There has traditionally been no "final examination" in writing courses at the University of Hawaii, and students are not required to reserve time for such an exam. We could not change this policy on short notice; aware of the traditional "no exam" policy, students had already made other plans for the "final examination" period that might otherwise have been the logical time for our posttest. And, even if it had been possible to use this time slot, there would have been objections to doing so: many instructors expressed the opinion that, if we gave the posttest in this time period, students would certainly believe--no matter what we told them--that the test was really to be used as a "final examination," would largely determine their grades for the course. We explored the possibility of scheduling a two-hour posttest at a time outside of class hours--possibly a Friday afternoon or a Saturday morning. We knew it would be difficult to find a suitable time, for many of our students work at outside jobs. But we thought we might solve that problem by scheduling the exam at several times, giving the student a choice of times. We soon found out that it was administratively impossible to require students to attend at <u>any</u> hour other than the already scheduled class hours. We could announce that the test would be given at certain hours, but we could not <u>require</u> students to attend; their attendance would have to be voluntary. And we knew that the results of our study would be completely unreliable if only volunteers took the posttest. So we were forced to use regular class hours for the posttest. We explored the possibilityof using two 50-minute class periods to make the total time allotted comparable to that allotted for the placement exam. But there were objections from instructors who had their class work schedules carefully planned and could not spare that much class time. And, besides, it was obvious that to use a two-period posttest would create conditions very different from those of the placement test anyway; an exam taken in two parts, with

considerable time in between in which to talk to other students about the exam, would obviously not yield the same results as a two-hour exam taken in one sitting. So we were forced to settle for a one-period, 50-minute posttest. We asked all instructors of English 100 to reserve 50 minutes of a class period near the end of the semester (specifically, April 25 for Monday-Wednesday-Friday classes, 50 minutes of the April 26 class period for Tuesday-Thursday classes) for the writing of an end-of-the-course essay. The students were told that these essays were required for a study we were conducting and that their performance on these essays would not affect their grades in the course.

The compromise we had to make with regard to the time allotted for the posttest necessitated a compromise in the content of the test as well. We decided that we had to combine into one question the two types of questions we had asked on the placement exam. We again gave the student two passages on the same subject (this time, Governor Ariyoshi's proposal for limiting immigration to Hawaii by reinstituting a residency requirement for government employment and for welfare benefits). And we again asked for an essay that compared and contrasted the two views. But this time we did not give the student a choice of new subjects to serve as the basis for a generalization that he would support with evidence from his own experience; instead, we asked the student to add to his comparison-contrast essay an expression of his own opinion on Ariyoshi's proposal. We realized that this would restrict the student's opportunity to demonstrate his rhetorical skills, for it would allow no choice of subject matter and it would make the "creative" part of the writing task simply a brief addendum to the essay that demonstrated reading comprehension and basic organizational skills. But the time limitation seemed to give us no choice.

We went through the end-of-the-course examinations and eliminated the papers written by students who had not taken the placement examination (these

students were upperclassmen who, for one reason or another, had not taken the freshman composition course when they were freshmen; they had been permitted to enroll in Group I sections). We had planned to take only a randomly selected sample of the end-of-the-course examinations to compare with the placement exams. But we were pleased to find that it was possible, within our time and budget limitations, to read all of the end-of-the-course exams written by students who had taken both the pretest and the posttest.

In reading the end-of-the-course exams, we used the same methods we had used in reading the placement exams. In fact, in an attempt to make the ratings as reliable as possible, we used the same readers: only faculty members who had participated in reading the placement exams were eligible to read the end-of-the-semester exams.

In addition to giving the same kind of holistic rating we had given on the placement exams, we attempted to use a second method of evaluation for the posttest: we rated some of the papers by assigning points for organization, wording, mechanics, etc., using as a checklist the Guidelines for Expository Writing worked out by the Composition Committee the year before. We found, however, that the reliability of the ratings given by this method was much lower than the reliability of the holistic ratings (i.e., there was greater disagreement among readers), and we decided that the holistic rating, which had proved to be very reliable, would give us all the information we needed.

After we had finished rating the posttests, we looked up the ratings the same students had been given on the pretest. We also compiled data on the student's sex and SAT scores, the type of grouping in the class the student had been in, the size of the class, and the instructor of the class.

We wanted to get subjective evaluations of the effects of our experiment in homogeneous grouping to supplement our statistical data. So, near the end of the semester we distributed two questionnaires: one to all the students in the homogeneous sections, another to all faculty members who had taught the

homogeneous sections. By tabulating and analyzing the answers given, we obtained some valuable information that could not have been obtained from the statistical data alone.

Finally, we looked into an effect of the homogeneous grouping that had no relationship to improvement of writing skills: we compared the drop-out rate in 1976-77, the year of our experiment in homogeneous grouping, with the drop-out rate in the preceding year.

RESULTS AND DISCUSSION

Statistical Data

On the basis of the placement examination, we divided the incoming freshmen as follows:

Total No. of Students Tested	Exempted from Taking Eng. 100	Placed in Group I	Placed in Group II	Placed in Group III	Placed in Group IV
2,212	44 (2%)	1,004 (45%)	498 (23%)	325 (15%)	341 (15%)

This distribution is not in the least startling; it is, in fact, very closely in line with national averages. The number of students exempted plus the number given a Group I rating equals 47% of the total; the number placed in Groups II, III, and IV equals 53% of the total. In other words, a little less than half of our students were judged to be ready for college-level writing; a little over half of our students were judged to need some degree of extra help to achieve a level of writing adequate for college-level work. These results are roughly in line with experience at other colleges and universities; for example, at the University of California at Berkeley, which admits only students who are in the top 12 1/2% of their high school classes, almost half of the students admitted in the fall of 1976 failed to pass the writing exam and were required to take the "Bonehead English" course (which is, incidentally, now provided as a part of the regular college offerings; the extra fee that used to be charged for the course has been abolished).

Our first task in beginning our study of the effects of homogeneous grouping was to determine the reliability of the ratings on the original placement exams. The number of readers involved in rating each exam varied: while approximately half of the exams had been assigned a final rating on the basis of three agreeing readers' ratings, other exams had required up to seven readings before three readers

awarded the same rating. This made it impossible for us to compute the usual correlation score for reader-reliability. Yet some measure of reliability seemed necessary. Some students had challenged the rating they received on their placement exams, pointing out that the exam had been given at an inopportune time for them. (Our decision to give the initial state-wide placement exam in area high schools whenever possible had made it necessary to give the exam while the high schools were still in session. Unfortunately, for some students the exam was given on the morning following their high school graduation, when they felt they were not in the best condition for writing.) Further, we knew that writing samples elicited during a single two-hour period might not be representative of the students' writing over a longer time period.

We finally decided to poll all teachers involved in the composition program, many of whom had been involved in rating the placement exams, asking them hypothetically to "place" their students on the basis of writing done for their courses during the first six weeks of the semester. We found a surprising degree of agreement: at each of the four homogeneous levels, teachers indicated that an average of 79% to 81% of the students had been placed at the correct level. There might have been a certain amount of self-fulfilling prophecy involved in these ratings. However, teachers with homogeneous classes seemed particularly sensitive to students who were either unable or more than able to keep up with their assignments. So we were quite confident of the accuracy of the instructors' ratings.

No similar reliability check was attempted with the end-of-course exam: instructors' ratings at the end of the semester might have been self-serving. But, since the end-of-course exam was rated only by readers who had previously rated the placement exams and the rating criteria were the same for both exams, we assumed a comparable degree of reliability.

Complete sets of data--placement exam scores, end-of-course exam scores, and SAT scores--were available for 424 students. SAT scores were missing for

only four students who had taken the end-of-course exam. So we based our entire statistical analysis on the 424 full sets of scores, and assumed that these scores were more or less representative of all the freshmen enrolled in expository writing courses.

The dependent measure used in the analysis of variance was the end-of-course exam score. There were three independent variables: type of grouping (homogeneous, heterogeneous, and "mixed"--the "mixed" group consisting of sections containing freshmen with placement exam scores of I and upperclassmen who had not taken a placement exam); level, as determined by the ratings on the placement exam (I-IV; the term "level" will be used from here on instead of "group" to avoid confusion with the "type of grouping" variable mentioned above); and class size (fewer than 15 students; 15 to 20; more than 20). SAT verbal and quantitative scores were used as covariates, since earlier analysis had revealed that SAT scores were significantly correlated with the end-of-course exam scores (with SAT verbal, $r = .35$; with SAT quantitative, $r = .19$).

The analysis of variance will be more meaningful if we first look at the actual mean scores attained by students in the various categories. Table 1 presents end-of-course exam score averages broken down by placement level and homogeneous/heterogeneous/mixed grouping. It also indicates the number of students in each of the categories.

TABLE 1

END-OF-COURSE EXAM AVERAGES* BROKEN DOWN BY PLACEMENT LEVEL AND GROUP, ALONG WITH NUMBER OF STUDENTS IN EACH CATEGORY

	Homogeneous	Heterogeneous	Mixed
Level I	3.565 (n=69)	3.897 (n=39)	3.706 (n=69)
Level II	3.525 (n=99)	3.346 (n=26)	
Level III	3.275 (n=51)	3.214 (n=14)	
Level IV	2.790 (n=43)	2.643 (n=14)	

*End-of-course exam ratings were converted to numbers by using the following scale: Exempt with Credit = 6, Exempt without Credit = 5, Level I = 4, Level II = 3, Level III = 2, Level IV = 1.

It is clear that, except at Level I, students in the homogeneous groups scored higher on the end-of-course exam than students in the heterogeneous groups. But the analysis of variance indicated that this difference was not great enough to be statistically significant ($F=0.10$, $df=2,378$).

It is also clear that the end-of-course exam scores were different at the four placement levels; students at the various placement levels did not attain comparable scores. And the analysis of variance indicated that this difference was great enough to be statistically significant ($F=9.51$, $df=3,378$, $p > .0001$).

Probably the most striking fact revealed by Table 1 is that the average score of 86% of the students was higher than 3.0. Given the nature of the end-of-course exam, this indicates that most students, at the end of the expository writing course, were communicating adequately, if not with full precision. Even those 14% of the students whose average scores were lower than 3.0 had shown a high degree of improvement--in fact, the greatest improvement of all--if we assume that their score on the end-of-course exam would have been 1.0 had the exam been given as a placement exam.

Perhaps the major question raised by these exam averages has to do with scores at Level I. The scores are lower than we would have expected had the exam been used as a placement exam. To a certain extent, these lowered scores can be attributed to the well-documented statistical phenomenon called "regression toward the mean." But the teacher ratings indicated that 84% of the students enrolled in Level I sections had been correctly placed via the placement exam. Are we to assume that some students who had scored high on the placement exam actually regressed as a result of having taken a course in expository writing?

It is here that we must consider carefully the essential difference between the placement exam and the end-of-course exam. The placement exam elicited two writing samples: the first called for evidence of understanding, organizational ability, and effective communication; the second called for similar evidence,

along with signs of ingenuity and creative exposition. The end-of-course exam was administered under different circumstances and allowed students less than one-half the amount of time they had had to write the placement exam. Its scope was necessarily limited to seeking evidence of basic expository writing skills. It is quite likely that many students earned high scores on the placement exam largely because of their creativity in exposition, which could outweigh minor deficiencies in writing skills. It is also probable that teachers with homogeneous Level-I sections encouraged the development of various rhetorical skills while assuming adequate facility with basic writing skills. Since the end-of-course exam was not designed to elicit evidence of sophisticated rhetorical skills, students who had spent time learning such skills did not really have a chance on the exam to demonstrate what they had learned.

This interpretation is supported by the differences in scores at Level I across the three experimental groups. Students in the heterogeneous groups, in which teachers had to emphasize basic skills to accommodate the needs of their less proficient students, scored highest on the end-of-course exam. Students in the mixed groups, in which teachers had to attend to the needs of their upperclass, non-placement-tested students, scored second best. Students in the homogeneous groups, where teachers assumed mastery of basic writing skills, obtained scores which indicated adequate development of basic skills, but they did not have a chance on this exam to demonstrate the more sophisticated skills they had probably acquired.

This particular line of reasoning is helpful in another way. It might be tempting, when looking at the raw mean scores presented in Table 1, to conclude that students matriculating with greater writing skills are likely to do better if placed in a heterogeneous group. But when we consider the restricted nature of the end-of-course exam, it is clear that our data cannot support such a conclusion.

On the basis of statistical analysis, then, we cannot conclude that either homogeneous or heterogeneous grouping is definitely more effective in teaching basic writing skills. The only conclusions we can draw are that the great majority of students in both homogeneous and heterogeneous sections of our composition course do improve in the basic writing skills measured by both the placement exam and the end-of-the-course exam and that students who enter the course with serious deficiencies in these basic writing skills improve more dramatically than those who are initially more proficient.

In an attempt to make sure that we did not attribute to homogeneous grouping any effects that should be attributed to other factors instead, we included, in our statistical data, information about the size of the class in which each student was enrolled. The analysis of variance revealed no significant interaction between grouping and class size ($F=1.82$, $df=4,378$).

We found that students in small classes generally showed greater improvement than students in larger classes. The analysis of variance did not indicate that this difference was statistically significant ($F=1.67$, $df=2,378$). Nevertheless, Table 2 reveals a trend, to which students at Level III were an exception: students in classes of 15 or fewer students improved in basic writing skills more markedly than did students in larger classes.

TABLE 2

END-OF-COURSE EXAM AVERAGES BROKEN DOWN BY PLACEMENT LEVEL AND CLASS SIZE, ALONG WITH NUMBER OF STUDENTS IN EACH CATEGORY

	Small Class (15 or fewer students)	Medium-Size Class (16-20 students)	Large Class (more than 20 students)
Level I	4.125 (n=16)	3.721 (n=61)	3.606 (n=100)
Level II	3.700 (n=10)	3.469 (n=32)	3.470 (n=83)
Level III	3.174 (n=23)	3.167 (n=12)	3.356 (n=30)
Level IV	3.000 (n=18)	2.737 (n=19)	2.550 (n=20)

The existence of such a trend is supported by the adjusted mean scores for class size yielded by the analysis of variance. When type of grouping, placement level, and SAT scores are held equal, we find that the mean end-of-course exam score for students in the small classes is 3.60 while the mean scores for students in the medium-sized and large classes are 3.42 and 3.40 respectively.

We cannot make much of these differences at this time. Both raw and adjusted mean scores suggest that limiting writing classes to 15 or fewer students enables students to make greater progress in developing basic writing skills. But a study in which students were more equitably distributed into classes of various sizes would yield a clearer indication of effects due to class size.

We obtained data on one possible effect of homogeneous grouping that had no relationship to improvement of writing skills: the effect of homogeneous grouping on the drop-out rate in classes. Table 3 shows the number and percentage of students who withdrew from the freshman composition classes in 1976-77, the year of our experiment, as compared with the preceding year, when all the classes were heterogeneous.

TABLE 3

WITHDRAWALS FROM ENGLISH 100 CLASSES, 1975-76 AND 1976-77

	Total Number of Students	Number Who Withdrew	Percentage Who Withdrew
Fall 1975	1335	137	10.2%
Fall 1976	1482	128	8.6%
Spring 1976	1089	185	16.9%
Spring 1977	937	96	10.2%

The data seem to show that homogeneous grouping lowered the drop-out rate in 1976-77--a fact that should be of some significance, for to reduce the number of withdrawals is to reduce costs. These results are not surprising: since

homogeneous grouping narrows the range in the skills of the students in the classes, it can be expected to reduce the number of students who find their particular class too difficult or too easy and, therefore, to reduce the number of students who withdraw from the classes. But, on the basis of a single year's experience with homogeneous grouping, we cannot assert that the reduction in the withdrawal rate was due entirely to homogeneous grouping; there may have been other factors involved. And, when we looked at the figures for Spring 1977 alone and compared the drop-out rate in the homogeneous sections with that in the heterogeneous control sections, we found that the withdrawal rate in the homogeneous sections was somewhat higher than that in the heterogeneous sections (10.8% as compared with 6.6%). These figures probably should not be given much significance, however, because the average size of classes was somewhat smaller in the heterogeneous sections; it seems quite possible that the advantage of smaller class size in discouraging withdrawals canceled out the disadvantage of heterogeneous grouping in this respect. Certainly, in order to rule out this possibility, classes of various sizes would have to be more equitably distributed than they were in our study. At present all we can say is that our data suggest that homogeneous grouping reduces the drop-out rate in classes. Further study of this relationship would seem to be justified.

Reactions of Students

A total of 400 students from 31 classes in which some homogeneous grouping had been attempted (including "mixed" classes containing untested upperclassmen as well as Level-I freshmen) responded to a single-item questionnaire distributed at the end of the semester. Students were asked to rate the helpfulness of having been placed in a class with other students who were at approximately the same level of writing proficiency; they were given time and space to add comments. Responses at each placement level are summarized in Table 4.

TABLE 4

STUDENT RATING, BY LEVEL, OF EFFECTIVENESS OF HOMOGENEOUS GROUPING

	Extremely Helpful	Somewhat Helpful	Inconsequential	Somewhat Hindering	Greatly Hindering
Level I*	33%	37%	28%	2%	0%
Level II	29%	54%	16%	1%	0%
Level III	22%	36%	36%	6%	0%
Level IV	55%	34%	7%	4%	0%

It is clear that the majority of students responding found homogeneous grouping "helpful"; at Level IV, a majority of students rated homogeneous grouping "extremely helpful." Responses at Level I were contaminated--that is, they included responses from upperclassmen who had been grouped with Level-I freshmen even though no upperclassmen had taken a placement exam. On the basis of the extended written comments we received, however, we can assume that if we had given the Level-I questionnaires only to freshmen who had experienced "uncontaminated" homogeneous grouping, their responses would have approximated those of students at Level IV.

We were not surprised that the strongest positive reactions to homogeneous grouping came from students at the highest and lowest levels. Written comments from students at these two levels were typically opposite: "We didn't have to waste time on basics because we all were pretty sure of the basics" (Level I) vs. "The teacher knew our problems from the beginning--he didn't have to waste time finding out" (Level IV); "It was challenging to be in a class where the comments from almost all of your classmates were useful" vs. "It was nice--no need to feel inferior to the smart asses!" Other comments from students at these two levels noted the general helpfulness of class sessions ("The teacher gave attention to my kind of problems") and the relevance of class discussions.

*Includes responses of untested upperclassmen placed with Level-I freshmen.

Finally, some Level-I students noted that they had put more effort into their writing because of the level of their "competition."

A number of comments from students at all levels had to do with course grades. Several students at Level I, instead of responding positively to the challenge of competition, expressed concern about the possibility that students at lower levels were getting comparable grades for less, or lower-quality, work. And some Level-IV students felt they might be getting higher grades than they would otherwise have received because they were not competing with more proficient writers.

Grading was, and continues to be, a problem. While we had discussed "general standards" with faculty members teaching the expository writing courses and had asked each one to remember that grades in homogeneous sections should not be expected to exhibit the "curve" typical of most heterogeneous sections, we were unable to impose strictly comparable grading standards: several faculty members indicated resentment of any attempt to curb their grade-assigning prerogatives. Thus, there undoubtedly were some cases in which unequal grades were assigned to equivalent efforts. One way out of this would be to base all course grades on end-of-course student essays read and graded by several faculty members, a procedure that has been advocated by some leaders in the field of research in composition.[1] However, we would undoubtedly have to overcome opposition from students as well as faculty members before we could adopt such a procedure.

Some negative comments from students at the middle levels were based on the fallibility of the original placement exam: they felt they had deserved higher or lower placement. "No single exam given when people are thinking about

---

[1] See, for example, Paul B. Diederich, "Comparative Preparation and Rating of Essay Tests," English Journal, 54:4 (April 1967), 573-584.

high school graduation should determine where you are placed" was a typical comment. And, until someone comes up with an infallible placement instrument, this argument will have merit, at least for about 20% of the students.

Another group of negative comments was based on the limited number of sections available to students because they had to register in sections designated as belonging to the level in which they had been placed. However, the apparent advantages of grouping made at least one convert: a Level-II student wrote "I never would have signed up for a 7:30 class if I had been given another choice [he or she had found all other Level-II sections closed by the time he or she arrived at registration]. But I think the advantages made getting up so early worth it."

Finally, upperclassmen enrolled with Level-I freshmen seemed generally less enthusiastic about homogeneous grouping because they felt that, as one put it, "Upperclass students should not have to compete with just the smart freshmen." Several upperclassmen felt that they too should have been given a placement exam. We agree with them, and would institute placement exams for all students if homogeneous grouping were to become standard.

In summary, most students perceived homogeneous grouping as helpful. Some based their perceptions on what they saw as advantages for learning; others based their perceptions on assumptions about grading procedures.

Reactions of Teachers

Thirty-nine teachers from the first semester, and twenty-five from the second, responded to our request for end-of-semester evaluations of the effects of homogeneous grouping.

Although we were primarily interested in the extended comments made by the teachers, we did tabulate responses to a question about the effects of homogeneous grouping on teaching. These responses are summarized in Table 5. There were no significant differences among teachers at different placement levels.

TABLE 5

TEACHER RATING, BY SEMESTER, OF RELATIONSHIP BETWEEN HOMOGENEOUS GROUPING AND TEACHING EFFECTIVENESS

	Made my teaching more effective	Made no difference in my teaching effectiveness	Made my teaching less effective	No basis for comparison
SEMESTER I	72%	15%	3%	10%
SEMESTER II	64%	15%	8%	12%

The majority of teachers felt that homogeneous grouping made their teaching more effective. Many commented that class time could be used more effectively when students had similar strengths and deficiencies ("I found that students on approximately the same level responded to each other, and to each other's writing, more readily"). Most noted that, in a homogeneous class, their class presentations seemed more relevant for more people. Many said that homogeneous grouping made their preparations easier, led to fewer worries over whom they were "getting through to" in their presentations, classroom activities, and writing assignments.

The greatest enthusiasm for homogeneous grouping seemed to come from teachers who worked exclusively with Level-I students. This enthusiasm was reflected in comments from a couple of teachers who taught classes made up of both Level-I freshmen and untested upperclassmen: they felt their freshmen were held back by the unpreparedness of some of their "upper" classmates. (Another teacher, however, said that the upperclassmen were the more talkative, and contributed to liveliness in the classroom.) One teacher noted a drawback she found with Level-I students only at the beginning of the first semester: "There were some snobs who felt that the placement exam had once again proved their superiority; I don't know why, but they gave a bad tone to the class, and seemed to think that there was nothing for them to learn." The reverse of this was evidently true in some of the lower-level classes, where teachers had to spend time during early class

sessions of the first semester assuring students that they weren't regarded as failures. However, no one made these comments about second-semester lower-level classes. Perhaps the early disappearance of the feeling that there was a stigma attached to being assigned to a low-level class was due to the fact that we were very careful not to describe our system as "ability grouping." We always spoke of it as a grouping according to kinds ofwriting problems the students had; we made it very clear that we considered the difference in the groups as a difference in preparation for college-level writing and not as a difference in intelligence. Also, it soon became known that the students in Level IV were working in very small groups--often even individually--with tutors and graduate assistants, and the idea of getting this kind of special help was very appealing to many students; we had many more requests for transfer from a higher-level section to a Level-IV section than requests for transfer from a Level-IV section to a higher-level section.

Teachers who indicated that homogeneous grouping had made no difference in, or had had a negative influence on, their teaching effectiveness added other useful comments. Typical of the former group: "I'd like to think that my teaching is always effective. But with grouping I find teaching easier. Preparation and class sessions and student appointments are all simpler because problems are similar for most students." One of the teachers who felt that homogeneous grouping had negative effects said that he preferred teaching a class with a wider range of ideas and writing styles. Another, teaching at Level IV, said "I found my IV's very passive. It was like storming a bastion of marshmallows."

Another measure of teacher response was not in written words. When it became known that there were no funds to administer a writing placement exam for next fall's incoming freshmen, teachers of composition generally reacted with

dismay. The dismay did not seem premature. It was not based on false assumptions that students in homogeneous classes learned better than did students in ungrouped classes; everyone knew that the results were not yet in. Instead, the dismay seemed to result from a sense that teachers had not yet really had enough of a chance to adjust to the advantages presented by grouped classes. "I was just beginning to see the possibilities of a grouped class, after a semester of trying to adjust my old syllabus to a Level-III class," said one teacher. "I wish I had another year to get my new act together. Now I guess I'll have to shape up that old syllabus one more time."

SUMMARY AND RECOMMENDATIONS

The primary question we were attempting to answer in this study was "Does homogeneous grouping in basic expository writing classes make teaching more effective?" The answer was not a clear, emphatic "Yes" or "No."

The statistical data were not conclusive. At all placement levels except Level I, students in the homogeneous groups scored higher on the end-of-course exam than students in the heterogeneous groups. It might be tempting to conclude that homogeneous grouping is more beneficial than heterogeneous grouping for students at the lower levels of proficiency but less beneficial for students at the highest level. But this conclusion cannot be supported by our data because of the compromises we were forced to make in designing the end-of-course exam; as explained in earlier sections of this report, the students at Level I might not have been an exception to the general rule if our end-of-course exam had given them greater opportunity to demonstrate the more sophisticated skills they had probably acquired. And, in any case, the differences between the end-of-course exam scores of the heterogeneous groups and the homogeneous groups were not great enough to be statistically significant.

The subjective evaluations of the effects of homogeneous grouping--from both teachers and students--provided a much clearer, more emphatic anwer to our primary question than the statistical data did. The great majority of both teachers and students were convinced that homogeneous grouping <u>did</u> make teaching in the basic expository writing courses more effective, <u>did</u> lead to greater improvement in the students' writing proficiency.

The results of our study are not conclusive enough to justify a demand for immediate adoption of homogeneous grouping in our basic expository writing classes. But they certainly do justify further research on the subject.

If we were to repeat our experiment, we would, of course, be able to re-

fine our procedures; the experience we gained in our first experiment would enable us to correct our mistakes. The most important change that would be necessary would be a change in the conditions and content of the posttest: before we began a new experiment, we would have to make provisions for a posttest that would be closely comparable to the pretest in all respects and would provide opportunity for students to demonstrate the variety of writing skills taught at all levels.

The possibility of conducting an experiment in homogeneous grouping without going to all the expense and work of giving a placement test of our own has been opened up by the College Entrance Examination Board's recent decision to incorporate an essay test in the Scholastic Aptitude Test that will be given in December this coming year. We have in the past not regarded a student's SAT verbal score--based as it is on an entirely objective, multiple-choice exam--as a reliable enough index of writing ability to use as a basis for grouping students; we do not believe that the fact that a student is able to identify a grammatically correct sentence necessarily means that he is also able to construct a series of logically related, coherent paragraphs. And we are not alone in this; the National Council of Teachers of English (NCTE) has issued a strong statement warning against the use of the SAT verbal score alone for placement purposes. But the contemplated change in the SAT may make the SAT score more useful to us. The CEEB's present plans call for only a 20-minute essay. And the essay will be incorporated in only the December test this coming year. But the possibility that a SAT verbal score that is based at least in part on the writing of an essay will be usable as a placement score is a possibility that we intend to investigate very carefully.

Our study of the effects of homogeneous grouping led us into another area that definitely seems to warrant further investigation: the effects of class

size. We found that students in classes of 15 or fewer students generally showed greater improvement in writing skills than students in larger classes. The differences were not great enough to be statistically significant. But our data suggest that limiting writing classes to 15 or fewer students would enable students to make greater progress whether the students were grouped homogeneously or not. We did not intend at the outset of this study to investigate the effects of class size, and therefore we made no attempt to distribute our students systematically into classes of various sizes. But the need to isolate the effects of homogeneous grouping from the effects of other factors forced us to take class size into consideration. And our findings certainly seem to justify a study in which students would be equitably distributed into classes of various sizes.

We did intend, at the outset of the study, to attempt to answer the questions "Do our students really show improvement in writing proficiency as a result of taking English 100?" and "Do our students meet at least minimum standards of proficiency at the end of the course?" We were pleased to find that our data answered both these questions in the affirmative. We had been prepared for a less flattering outcome. We were aware that "there has been very little evidence from research to support the contention that instruction in composition does have a significant favorable effect on student writing."[1] And we knew we could not, in this study, duplicate the type of test that had demonstrated improvement most clearly: the writing of researched essays on topics of the students' own choosing for both the pretest and the posttest.[2] So we were pleased to find that the great majority of our students scored higher

---

1. Sara E. Sanders and John H. Littlefield, "Perhaps Test Essays Can Reflect Significant Improvement in Freshman Composition: Report on a Successful Attempt," Research in the Teaching of English, 9 (Fall, 1975), 146.

2. Ibid., 145-153.

on the end-of-course exam than they had on the placement exam and in fact, by the end of the course, were communicating at a level that we regarded as meeting at least minimal standards of proficiency.

We certainly are not satisfied with the level of proficiency attained by our students, however. The high reliability of the ratings of the readers of the placement exams indicates that we have established goals. But we are aware that we need to work on clarifying our goals further for ourselves and for our students and we are aware that we need to improve classroom techniques to help students to reach those goals. And this will require further research. This study has been essentially a pilot study; perhaps its greatest value has been that it has suggested important areas which demand further analysis, experimentation, and evaluation.